

This is a repository copy of *Comparison of two independent systematic reviews of trials of recombinant human bone morphogenetic protein-2 (rhBMP-2) : The Yale Open Data Access Medtronic Project*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/115322/>

Version: Published Version

Article:

Low, Jeffrey, Ross, Joseph S., Ritchie, Jessica D. et al. (6 more authors) (2017)
Comparison of two independent systematic reviews of trials of recombinant human bone morphogenetic protein-2 (rhBMP-2) : The Yale Open Data Access Medtronic Project. Systematic Reviews. 28. ISSN 2046-4053

<https://doi.org/10.1186/s13643-017-0422-x>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

RESEARCH

Open Access



Comparison of two independent systematic reviews of trials of recombinant human bone morphogenetic protein-2 (rhBMP-2): the Yale Open Data Access Medtronic Project

Jeffrey Low¹, Joseph S. Ross^{2,3,4,5}, Jessica D. Ritchie², Cary P. Gross^{3,4}, Richard Lehman⁶, Haiqun Lin^{2,7}, Rongwei Fu⁸, Lesley A. Stewart⁹ and Harlan M. Krumholz^{2,4,5,10*}

Abstract

Background: It is uncertain whether the replication of systematic reviews, particularly those with the same objectives and resources, would employ similar methods and/or arrive at identical findings. We compared the results and conclusions of two concurrent systematic reviews undertaken by two independent research teams provided with the same objectives, resources, and individual participant-level data.

Methods: Two centers in the USA and UK were each provided with participant-level data on 17 multi-site clinical trials of recombinant human bone morphogenetic protein-2 (rhBMP-2). The teams were blinded to each other's methods and findings until after publication. We conducted a retrospective structured comparison of the results of the two systematic reviews. The main outcome measures included (1) trial inclusion criteria; (2) statistical methods; (3) summary efficacy and risk estimates; and (4) conclusions.

Results: The two research teams' meta-analyses inclusion criteria were broadly similar but differed slightly in trial inclusion and research methodology. They obtained similar results in summary estimates of most clinical outcomes and adverse events. Center A incorporated all trials into summary estimates of efficacy and harms, while Center B concentrated on analyses stratified by surgical approach. Center A found a statistically significant, but small, benefit whereas Center B reported no advantage. In the analysis of harms, neither showed an increased cancer risk at 48 months, although Center B reported a significant increase at 24 months. Conclusions reflected these differences in summary estimates of benefit balanced with small but potentially important risk of harm.

Conclusions: Two independent groups given the same research objectives, data, resources, funding, and time produced broad general agreement but differed in several areas. These differences, the importance of which is debatable, indicate the value of the availability of data to allow for more than a single approach and a single interpretation of the data.

Systematic review registration: PROSPERO CRD42012002040 and CRD42012001907.

Keywords: Data sharing, Data interpretation, Systematic review, Meta-analysis

* Correspondence: harlan.krumholz@yale.edu

²Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, 1 Church Street, Suite 200, New Haven, CT 06510, USA

⁴Robert Wood Johnson Foundation Clinical Scholars Program, P.O. Box 208088, New Haven, CT 06510, USA

Full list of author information is available at the end of the article



Background

Systematic reviews and meta-analyses [1] based on individual participant-level data (IPD) from randomized controlled trials (RCTs) are considered to provide the highest level of rigor for evaluating the evidence for a clinical question [2]. Such reviews offer the possibility of using hierarchical statistical techniques that better handle sources of heterogeneity, allow for sub-group analyses, and facilitate assessment of rare events. Previously, IPD meta-analyses have modified [3–8] or overturned [9] the results of previous meta-analyses based on the published literature alone.

Efficient and unbiased mechanisms to replicate research findings are essential for maintaining high levels of scientific credibility [10]. The premise of replication efforts is that different groups, employing rigorous methods, may take different approaches and come to different conclusions on a previously addressed question. Recent efforts to promote data sharing by the National Institutes of Health, [11, 12] the pharmaceutical industry, [13, 14] and partnerships between academia and industry [15, 16] have made replication an increasingly available mechanism to test the validity of clinical trial conclusions. This work is particularly important for systematic reviews and meta-analyses, which frequently form the basis of professional society and government guideline recommendations [17].

Previous studies have sought to determine whether systematic reviews are replicable, with new teams performing new searches, summaries, and analyses of the literature for a particular question. These studies, which compare systematic reviews of the published literature conducted at different time points, suggested that groups investigating the same research question may differ in their findings [18–21], though most often, these differences were attributed to search strategy [6, 19, 22–25]. However, it is uncertain if replication of meta-analyses, particularly those with the same research objectives, participant-level data, time, and funding, would employ the same analytic methods or arrive at identical findings. A thorough understanding of the reliability of meta-analysis requires an empiric assessment of how two distinct teams of investigators would employ meta-analytic techniques to address the same clinical question. Accordingly, we sought to determine if two independent centers, each of which were contracted to pursue identical research questions concurrently, with access to identical IPD, would employ identical methods in the areas of data use and statistical analysis and report identical, or at least consistent, results and conclusions.

Methods

Study design

We retrospectively compared the research methods and results of the final comprehensive publications of two meta-analyses performed in the context of full systematic

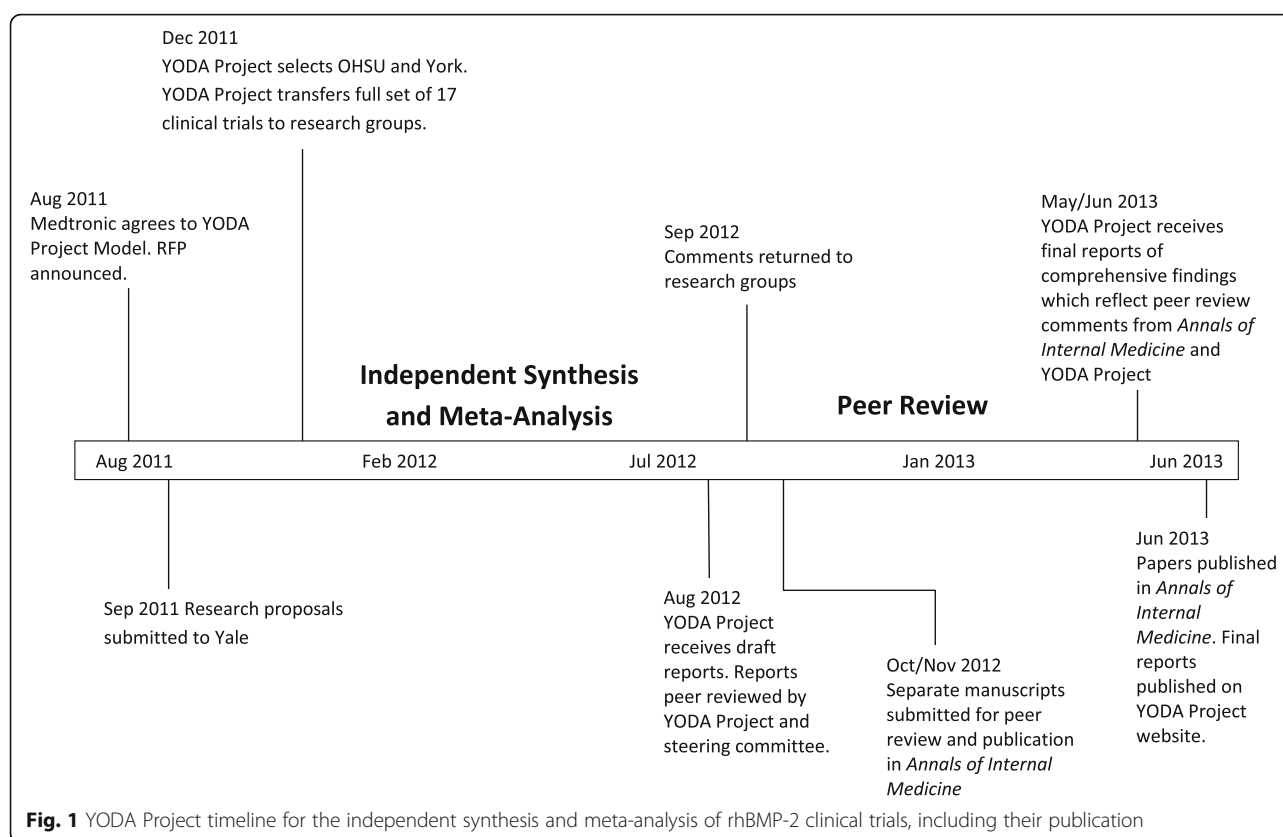
reviews of recombinant human bone morphogenetic protein-2 (rhBMP-2) prepared by two independent centers, Center A [26, 27] from the University of York and Center B [28, 29] from Oregon Health & Science University, and focused on (1) meta-analysis trial inclusion criteria; (2) statistical methods; (3) summary risk estimates; and (4) conclusions.

Trial inclusion criteria were defined as study characteristics necessary for inclusion in meta-analysis. We explicitly compared, for primary and secondary endpoint meta-analyses, as well as safety analyses, the trials used by both centers for each analysis. For methods, we compared centers' reported outcomes at various time points as well as statistical methods. We compared the centers' risk estimates for all primary outcomes for efficacy as well as safety at all time points. In consideration of these factors, we provide a subjective comparison of the overall conclusions drawn by each center.

Conducting the systematic reviews and IPD meta-analysis

Following controversy in the literature surrounding adverse events related to rhBMP-2 including cancer, in August 2011, Medtronic agreed to participate in the Yale University Open Data Access (YODA) Project model, which has been described previously (Fig. 1) [30]. Appendix 1 provides additional context on the particular clinical controversy covered by these reviews. Our analysis will focus on systematic review reproducibility rather than this particular clinical question which has already been well described in the literature. An open request for proposal was announced by the YODA Project to solicit applications from external investigators with preliminary research aims to study the safety and efficacy of rhBMP-2. The YODA Project selected research groups from Oregon Health & Science University (OHSU) and the University of York in the UK (York). These leading centers specialize in the conduct of systematic reviews and bring internationally recognized primary investigators who have made significant contributions to methodology development for organizations including the Cochrane Collaboration and the Agency for Healthcare Research and Quality (AHRQ). Based on feedback from OHSU and York, a set of reconciled aims were developed to ensure a common scope (Table 1) [31]. Each group independently developed its protocol for conducting the systematic review and deposited the full protocol with the YODA Project. Both groups registered short versions of their protocols without detailed methods for analysis on the PROSPERO registry of systematic reviews on February 23, 2012 (CRD42012002040 and CRD42012001907).

The YODA Project transferred the full set of Medtronic data relating to rhBMP-2 to the centers in early December 2011. This included full de-identified individual participant-



level data for 17 trials, consisting of 8 pilot studies, 8 pivotal RCTs, and 1 study terminated for commercial reasons. The total number of participants was 2091, consisting of 1077 rhBMP-2 recipients and 1014 control participants. Also included were protocols, data dictionaries, internal reports consisting of summaries of study data, and brief adverse event case histories. In addition, 1229 MedWatch adverse event reports submitted to the US Food and Drug Administration between July 2003 and July 2012 were provided.

Each center completed IPD meta-analyses on the effectiveness and harms of rhBMP-2 in the context of full systematic reviews. Each site was responsible for determining the appropriateness of conducting a systematic review as well as its methods and research questions within the scope of the specified research aims. The project was designed so the review groups would work in parallel and have no mutual communication about their approaches. Questions from the groups were communicated through the YODA Project review coordinator so that there was no direct communication between the groups and Medtronic.

Draft reports of comprehensive findings were received from both groups by the YODA Project in mid-August 2012. These reports were peer-reviewed by separate review teams consisting of members of the YODA Project and steering committee, which included clinical, statistical, and methodological experts, as well as by a

representative from Medtronic. A peer reviewer had access to only one of the two reports at any time before final publication, and there was no communication between the separate review teams. Comments were returned to the research groups in September 2012. The groups prepared separate manuscripts for submission for publication in the *Annals of Internal Medicine*. Final reports of comprehensive findings, which reflected peer review comments from the journal and from the YODA Project, were received in summer 2013. These comprehensive reports, which we review in this paper, were published on the YODA Project website congruently with the articles in the *Annals* on June 18, 2013. The data set has subsequently been made available to additional researchers through a request process [32]. The Human Investigation Committee at Yale University determined that this study is not considered to be Human Subjects Research and did not require further review.

Results

Meta-analysis inclusion criteria

Trial inclusion was largely similar with a primary difference of IPD obtained from a single published RCT. Both centers chose only to include RCTs of rhBMP-2 in spinal fusion in their meta-analysis, and both groups analyzed 11 of the RCTs. Center A obtained IPD from, and included an additional non-industry sponsored RCT by, Glassman

Table 1 Explicit research aims provided to the 2 independent centers by the YODA Project

Research aims
1. Identify all relevant studies, across all uses and sponsor (i.e., Medtronic sponsored, others).
2. Determine the questions that were addressed by these studies.
3. Evaluate the quality of the studies. Assess the risk of bias associated with the design, conduct, and reporting of each clinical study, including those identified via the systematic review and those provided by Medtronic, and, if present, how bias may have affected assessment of effectiveness and harms.
(a) Assessment of study design and conduct should include evaluation of internal validity, methods used to ascertain outcomes and other policies and procedures for data collection, as well as the integrity of case report form adjudication.
(b) Assessment of study reporting should include selective publication and selective reporting
(c) Summary of these findings should include:
i. What conclusions can be drawn by assessing the full body of data and what gaps in knowledge remain, taking into account results from the evaluation of quality and risk of bias
ii. An assessment of applicability of these studies
4. Conduct meta-analyses from studies identified via the systematic review, if appropriate, and using patient-level data, if possible. If not appropriate there should be another approach to summarizing the data. The analysis should consider the following:
a. For effectiveness, meta-analysis should consider patient-centered outcomes (i.e., quality of life and functional status), as well as surrogate outcomes (i.e., fusion as determined by radiography).
b. For safety, meta-analysis should include local effects such as inflammation, heterotopic bone formation, pain, osteolysis and instability, and downstream or systematic effects such as leg pain and weakness, retrograde ejaculations, and possible increased risk of cancer.

et al. [33] for its analysis of effectiveness but excluded it when looking at harms since events were reported differently and without information on when they occurred. Though Center B identified this study, it did not solicit IPD from its authors and was able to include only a qualitative analysis.

Research methodology

Research methodology differed primarily in the choice of stratification, with minor differences in the choice of statistical methods. For analyses of benefits, Center A included trials that compared rhBMP-2 with standard bone grafting techniques across all surgical approaches. As the primary analysis, Center A performed a standard two-stage meta-analysis along with a sub-group analysis that did not find evidence of differences between surgical approaches.

Center B stratified by surgical approach for effectiveness and most harms and determined that only two of the four surgical approaches (anterior lumbar interbody fusion (ALIF) and posterolateral fusion (PLF)), which were studied in multiple RCTs, provided adequate data

for meta-analysis. Center B employed a one-stage meta-analysis, using mixed effects regression models. The study comparing rhBMP-2 with lumbar disc prosthesis was included in the analysis of cancer and death, which was not stratified by surgical approach.

Both centers studied the same primary outcomes for effectiveness and reported them at the same time points of 6 weeks and 3, 6, 12, and 24 months after surgery (Table 2).

Table 2 Results from meta-analyses conducted independently by Centers A and B examining measures of efficacy and safety associated with rhBMP-2

Outcome	Center	Surgical approach	No. of studies (n)	Effect size (95% CI)	Treatment advantage
ODI	A	All	12 (1368)	(0-100) -3.48 (-6.47 to -0.49)	BMP
		ALIF	5 (423)	(0-50) -7.35 (-14.00 to -0.70)	BMP
		PLF	4 (650)	(0-50) -1.98 (-4.86 to 0.90)	Neither
SF-36 PCS	A	All	12 (1303)	(0-100) 1.93 (0.63 to 3.22)	BMP
		ALIF	5 (421)	(0-100) 3.68 (0.86 to 6.49)	BMP
		PLF	4 (644)	(0-100) 1.10 (-0.6 to 2.86)	Neither
Back pain	A	All	12 (1326)	(0-10) -1.58 (-2.65 to -0.51)	BMP
		ALIF	4 (409)	(0-20) -0.74 (1.49 to 0.00)	BMP
		PLF	4 (649)	(0-20) -0.31 (-0.76 to 0.15)	Neither
Leg pain	A	All	12 (1326)	(0-10) -0.59 (-1.27 to 0.09)	Neither
		ALIF	4 (409)	(0-20) -0.60 (-1.28 to 0.08)	Neither
		PLF	4 (648)	(0-20) -0.34 (-0.82 to 0.13)	Neither
Fusion	A	All	10 (1078)	RR 1.14 (1.03 to 1.25)	BMP
		ALIF	5 (416)	RR 1.05 (0.88 to 1.24)	Neither
		PLF	4 (637)	RR 1.16 (0.96 to 1.41)	Neither
Cancer	A	All up to 48 months	11 (1281)	RR 1.98 (0.86 to 4.54)	Neither
		All 24 months	5 (1450)	RR 3.45 (1.98 to 6.00)	Control
		All 48 months	4 (1183)	RR 1.82 (0.84 to 3.95)	Neither

CI Confidence Interval, ODI Oswestry disability index. Lower favors rhBMP-2. ALIF Anterior lumbar interbody fusion, PLF Posterolateral fusion, RR Relative risk, SF-36 PCS Short Form 36 Physical Component Score. Higher favors rhBMP-2. For back and leg pain, Center A used a 0-10 scale and Center B used a 0-20 scale. Lower favors rhBMP-2. Bolded values statistically significant

Similar outcomes were also reported between the centers for harms up to 4 weeks and then up to 24 months for general adverse events, and up to 48 months for cancer and death.

Neither group found evidence of an rhBMP-2 dose-response relationship or heterogeneity in groups that received high-dose forms of rhBMP-2, so all dose formulations were combined.

For harms, Center A chose to combine all trials using a generalized mixed effects model since specific adverse events were few at the trial level. Center B also used a generalized mixed effects model with stratification by surgical approach, except for cancer and death.

Summary results estimates

The groups obtained similar results in summary estimates of most clinical outcomes and adverse events, although there were notable differences. Center A found a statistically significant increase in fusion rate at 24 months (12% over controls) combining data across all surgical approaches. In contrast, Center B, reporting results for each surgical approach separately, did not find a significant increase in fusion at 24 months. For reducing back pain and overall disability, Center A found a statistically significant advantage for all time points from 6 months onwards when combining data from all approaches, with no statistically significant difference in the effectiveness of rhBMP-2 by surgical approach (Fig. 2). For Center B, differences for pain reduction were statistically significant

from 3 months onward for ALIF, but only at the 6-month time point for PLF.

Findings were not identical for cancer; Center B reported a statistically significant increased risk at 24 months with the use of rhBMP-2, and Center A did not report at 24 months. Neither group found a significantly increased risk of cancer associated with rhBMP-2 at 48 months (Fig. 2). Both groups reported similar but not identical findings for the frequency of regular adverse events.

Summary conclusions

Center A interpreted benefits to fusion and postoperative pain as “clinically insignificant” and increased cancer incidence as “inconclusive,” noting that “whether this increased risk is genuine is uncertain” (Table 3). Overall, by this analysis alone, rhBMP-2 seemed to offer improved rates of fusion with similar clinical outcomes compared with standard techniques at the expense of increased reports of back and leg pain in the early postoperative period.

In contrast to Center A’s report, Center B found “moderate-strength evidence of no consistent differences between rhBMP-2 and ICBG in...fusion rates.” In addition, it reported a statistically significant increase in cancer at the 24-month time point, while noting that “This finding should be interpreted with caution because cases were heterogeneous.” Overall conclusions from Center B seemed to indicate more strongly than those of Center A that rhBMP-2 had no additional clinical benefit. Center B reported that its “analysis underscores that more definitive evidence about harms was needed before rhBMP-2

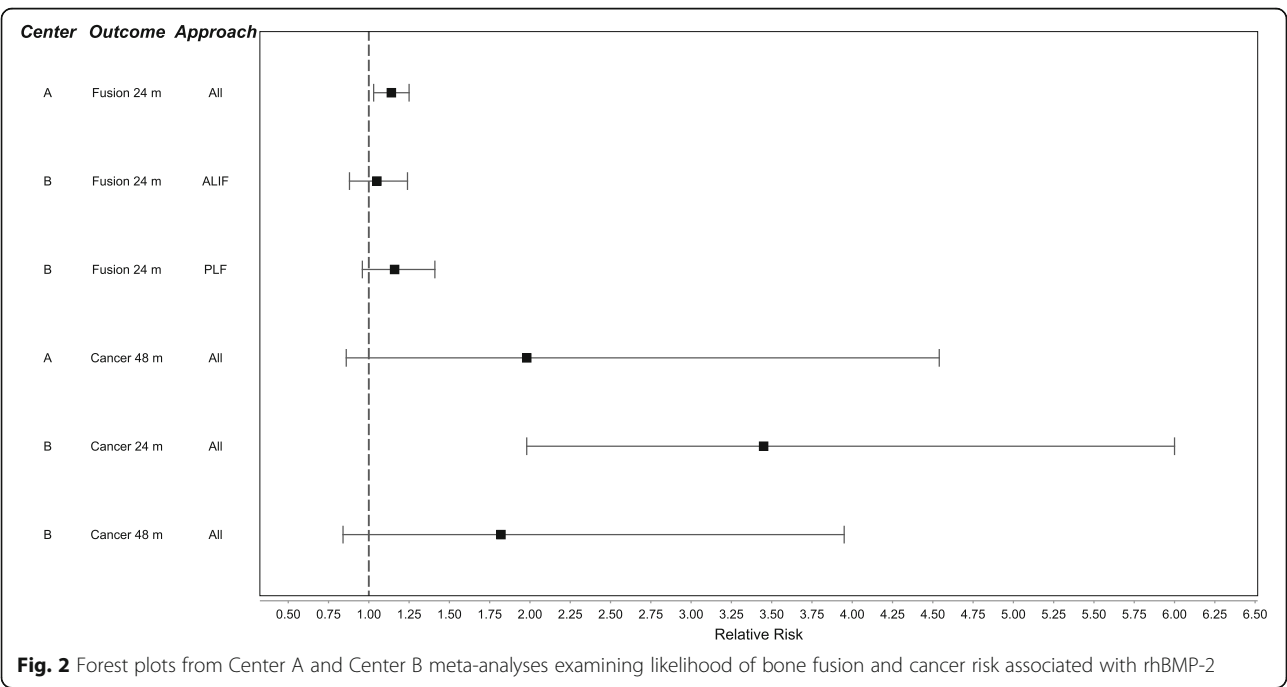


Fig. 2 Forest plots from Center A and Center B meta-analyses examining likelihood of bone fusion and cancer risk associated with rhBMP-2

Table 3 Conclusions made by Centers A and B after conducting independent summary reviews and meta-analysis of rhBMP-2 trials

Center	Conclusions
A	<p>"At 24 months, rhBMP-2 increases fusion rates, reduces pain by a clinically insignificant amount, and increases early postsurgical pain compared with ICBG. Evidence of increased cancer incidence is inconclusive."</p> <p>"The use of rhBMP-2 in spinal fusion surgery increases the likelihood of successful fusion at up to 24 months, but this does not seem to translate into a clinically significant reduction in pain. The small improvements in fusion and in the level of pain reduction, which manifest after 6 months, also seem to come at the expense of more frequent pain in the immediate postoperative period and, possibly, an increased number of cancer cases. We believe that it is important that clinicians explain these findings to patients so that they can make informed choices about the type of surgery they would prefer."</p>
B	<p>"In spinal fusion, rhBMP-2 has no proven clinical advantage over bone graft and may be associated with important harms, making it difficult to identify clear indications for rhBMP-2."</p> <p>In conclusion, we found...no evidence that rhBMP-2 is more effective than ICBG in spinal fusion, with some evidence of an association with important harms. More research is needed to provide more reliable estimates of risk for cancer and other adverse events and to identify patient populations in which use of rhBMP-2 may be beneficial, such as cases where use of bone graft alone is associated with a high risk for pseudarthrosis. On the basis of the currently available evidence, it is difficult to identify clear indications for rhBMP-2 in spinal fusion."</p>

became widely used" and that "On the basis of the currently available evidence, it is difficult to identify clear indications for rhBMP-2 in spinal fusion. This analysis shows almost no clinical benefit for the product while raising questions about the potential risk for cancer."

Discussion

In our study of two independent centers provided with identical objectives, data, resources, and time to conduct concurrent meta-analyses, we found that the centers did not report identical methods, results, and interpretations. In addition, the potential benefit of additional analyses of the same data was not limited solely to increasing confidence through replication. Separate analyses revealed nuances of differences, with potential interpretations for clinical management, which could be produced from the same data set using valid methods. These findings, even though largely similar, support the case for greater sharing and access to clinical data as a way to maximize public dialogue about the meaning of the data and to ensure that a single interpretation does not lead people to believe there is no other possible approach.

The centers took different but methodically defensible approaches in their attempts to best represent the results of this data set in a relevant and valid way. Review methods differed based on data stratification and IPD obtained from an additional trial. One group chose to

combine data across all surgical approaches, finding little heterogeneity in trials by approach. The other group chose to stratify and analyze by surgical approach, forgoing increased statistical power in recognition of the real differences and adverse event concerns between different surgical approaches, and to present in a format perhaps more intuitive to spine surgeons. Study inclusion diverged, with one group obtaining IPD from an additional trial not funded by Medtronic and including it in the analysis of effectiveness. Even with the proliferation of standards in methodology, this demonstrates that we can expect some differences in how two similarly qualified groups might choose to conduct a complex systematic review. This diversity in methods has the potential to add to the depth of our understanding of a product and reinforces that additional value can be tapped from a data set with open access.

These differences in approach led to differences in summary estimates. In the case of the outcome of spinal fusion, this led to a difference that had statistical relevance even as the group discounted the clinical importance. Nevertheless, this finding could support the argument in the spine literature that the use of this product is warranted in certain indications and select cases where the risk of non-union is great and its consequences potentially disastrous [34, 35]. In contrast, this difference was no longer detectable when data were stratified by a surgical approach in the other review, and surgeons looking at these data alone might see fewer instances where this product would be beneficial. For estimates of cancer, there were slight differences in the time points reported. Center B showed a statistically significant increase in cancer at the 24-month time point but concurred that cancer was not significantly increased for longer follow-up. In both cases, the absolute risk of cancer was low, and the cancer types represented were heterogeneous.

In contrast to previous studies of concurrent meta-analyses in nutrition and endometrial cancer [7] and immunotherapy treatments for spontaneous abortions, this study found that concurrently conducted meta-analyses examining the same data arrived at conclusions that readers may or may not interpret similarly. Currently, it appears that nearly all meta-analyses are conducted by single groups, without replication. Additional analyses necessitate the sharing of data, and this in itself can bring important benefits. Information in the published literature is often incomplete. Data sharing has the potential to allow for a more complete picture of the benefits and harms of a treatment based on the totality of available evidence. Data are often collected or subsidized at the public expense and need to be made more widely available for the public benefit. Across a diverse array of fields, open access to data and the potential for reanalysis can, at the minimum, strengthen confidence in the findings of a

systematic review while offering the potential to add to or even alter the conclusions about an intervention.

While there are many benefits and arguments for greater data sharing, these benefits must be considered in light of potential downsides that might come with additional analyses based on the same data. In this project, we addressed industry concerns around spurious analysis and litigation, as well as biased and methodologically flawed studies which might unfairly taint a product. Academia too faces challenges around credit, bias, and the potential for conflicting messages to confound decision-making. Ultimately, we believe a process of frameworks, like the YODA Project, and norms can help manage these potential problems and unlock the benefits that come with greater sharing of data.

The generalizability of our findings to other settings is not known. However, the design of our approach should have made it more likely that the results would have been the same rather than different. The two groups were provided with the same data from all manufacturer-sponsored studies which, for rhBMP-2, represented the vast majority of high-quality studies on this product. Studies of the other questions could be limited by differences in search strategies and disagreement over key studies. The groups also received identical funding from an outside organization, and neither the groups nor the funders had any financial interest in this product.

Conclusions

Two independent and expert review groups that performed independent meta-analyses of rhBMP-2 came to broadly similar findings, though with some differences on the statistical significance of primary analyses of fusion and cancer. The clinical importance of the differences may be debatable, and even the authors of this article differed in their interpretations of the results and conclusions presented in these analyses. What is certain is that the methods and interpretations were not identical and had different points of emphasis. This underscores the importance of making data more openly available for the purpose of additional scientific inquiry to maximize the knowledge that can be extracted.

Appendix 1

rhBMP-2 background and summary findings

Clinical background: spinal fusion is a commonly performed surgery to correct spinal instability and commonly accompanies spinal decompressions among other uses. To improve the rate of fusion if local bone is insufficient, surgeons often use graft material. Autologous iliac crest bone graft (ICBG) is considering the gold standard for grafting but often involves harvesting bone from a separate site through a separate incision.

Recombinant bone morphogenetic protein-2 (rhBMP-2) is an orthobiologic that promotes bone formation and spinal fusion that is used as an alternative to ICBG. It is approved for anterior lumbar interbody fusion (ALIF) surgery but was widely used in other surgical approaches with off-label use estimated at 85% [36] and peak sales approach \$1 billion [37].

Adverse events attributed to rhBMP-2 began to be reported. In 2008, the US Food and Drug Administration (FDA) issued a warning about dangerous side effects associated with rhBMP-2 use in anterior cervical spine surgery including dysphagia, hematoma, seroma, swelling, and the need for intubation. Subsequent studies brought attention to unexpected adverse events in surgical approaches outside the approved anterior lumbar approach including radiculitis, vertebral body resorption, seroma and/or hematoma formation, and heterotopic ossification. Even in the approved ALIF approach, there were reported associated with osteolysis and retrograde ejaculation [38]. A review of publicly available data in 2011 sparked the controversy which led to this systematic review effort asserting an underreporting of adverse events including a risk of cancer [39].

Abbreviations

AHRQ: Agency for Healthcare Research and Quality; ALIF: Anterior lumbar interbody fusion; ICBG: Iliac crest bone graft; IPD: Individual participant-level data; OHSU: Oregon Health & Science University; PLF: Posterolateral fusion; RCT: Randomized controlled trials; rhBMP-2: Recombinant human bone morphogenetic protein-2; YODA: Yale University Open Data Access; York: University of York

Acknowledgements

The authors thank Mark Simmonds and Mark Rodgers (Team A, University of York) and Roger Chou and Marian McDonagh (Team B, Oregon Health and Science University) for commenting on the previous drafts of this paper.

Funding

Dr. Krumholz is supported by grant U01 HL105270-05 (Center for Cardiovascular Outcomes Research at Yale University) from the National Heart, Lung, and Blood Institute. Dr. Ross is supported by grant K08 AG032886 from the National Institute on Aging and by the American Federation for Aging Research through the Paul B. Beeson Career Development Award Program. The funders were not involved in the design of the study, conduct of the work, or the development and submittal of the work for publication.

Authors' contributions

HK conceived and initiated the collaborative project and is the guarantor. JL was responsible for the acquisition of the data and drafted the manuscript. JSR, JDR, CG, RL, HL, RF, and LS interpreted the data and revised the paper for important intellectual content. HK affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. All authors reviewed and approved the final manuscript.

Competing interests

Drs. Gross, Krumholz and Ross are funded by research agreements from Medtronic and from Johnson & Johnson (Janssen), through Yale University, to develop methods of clinical trial data sharing. Drs. Krumholz and Ross work under contract to the Centers for Medicare & Medicaid Services to develop and maintain performance measures and are the recipients of research support from the Food and Drug Administration, through Yale University, to develop methods for post-market surveillance of medical devices.

Dr. Krumholz chairs a cardiac scientific advisory board for UnitedHealth. Dr. Gross receives research funding from 21st Century Oncology. Dr. Fu previously received funding from the Yale Open Data Access (YODA) Project to carry out the analyses attributed to Team B in this paper. Dr. Stewart is employed as Director of the Centre for Reviews and Dissemination at the University of York and is in receipt of research funding to carry out health technology assessments including systematic reviews and meta-analyses for the National Institute for Health Research, and she previously received funding from the YODA project to carry out the analyses attributed to Team A in this paper; the Centre for Reviews and Dissemination has a policy not to conduct work for or on behalf of the pharmaceutical or medical devices industry. The other authors do not have competing interests to report.

Author details

¹Department of Internal Medicine, Yale School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA. ²Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, 1 Church Street, Suite 200, New Haven, CT 06510, USA. ³Section of General Internal Medicine, Department of Internal Medicine, Yale School of Medicine, P.O. Box 208093, New Haven, CT 06510, USA. ⁴Robert Wood Johnson Foundation Clinical Scholars Program, P.O. Box 208088, New Haven, CT 06510, USA. ⁵Department of Health Policy and Management, Yale School of Public Health, P.O. Box 208034, New Haven, CT 06520, USA. ⁶Department of Primary Care Health Sciences, Oxford University, 33 St Ebbses Street, Oxford OX1 1PU, UK. ⁷Department of Biostatistics, Yale School of Public Health, P.O. Box 208034, New Haven, CT 06520, USA. ⁸Department of Public Health and Preventive Medicine, Pacific Northwest Evidence-based Practice Center, Oregon Health & Science University, 3181 S.W. Sam Jackson Park Road, Portland, OR 97239-3098, USA. ⁹Centre for Reviews and Dissemination, University of York, Heslington, York YO10 5DD, UK. ¹⁰Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, P.O. Box 208093, New Haven, CT 06510, USA.

Received: 22 February 2016 Accepted: 25 January 2017

Published online: 15 February 2017

References

- Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*. 2005;294(2):218–28.
- Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof*. 2002;25(1):76–97.
- Wang JG, Staessen JA, Franklin SS, Fagard R, Gueyffier F. Systolic and diastolic blood pressure lowering as determinants of cardiovascular outcome. *Hypertension*. 2005;45(5):907–13.
- Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet*. 1993;341(8842):418–22.
- Riley RD, Lambert PC, Staessen JA, Wang J, Gueyffier F, Thijs L, Bouitrie F. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med*. 2008;27(11):1870–93.
- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*. 2010;340:c221.
- Jeng GT, Scott JR, Burmeister LF. A comparison of meta-analytic results using literature vs individual patient data. *Paternal cell immunization for recurrent miscarriage*. *JAMA*. 1995;274(10):830–6.
- Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Anti-lymphocyte antibody induction therapy study G: individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med*. 2002;21(3):371–87.
- McCormack K, Grant A, Scott N. Collaboration EUHT. Value of updating a systematic review in surgery using individual patient data. *Br J Surg*. 2004;91(4):495–9.
- Ioannidis JP. Why science is not necessarily self-correcting. *Perspect Psychol Sci*. 2012;7(6):645–54.
- Final NIH statement on sharing research data [http://grants.nih.gov/grants/policy/data_sharing/]. Accessed 8 Apr 2015.
- NOT-OD-15-019: NIH request for public comments on the draft NIH policy on dissemination of NIH-funded clinical trial information [<http://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-019.html>]. Accessed 8 Apr 2015.
- Data sharing commitments enhance research and scientific knowledge, advance patient care and improve public health [<http://www.phrma.org/Joint-EFPIA-PhRMA-Principles-for-Responsible-Clinical-Trial-Data-Sharing-Become-Effective-Today>]. Accessed 8 Apr 2015.
- Nisen P, Rockhold F. Access to patient-level data from GlaxoSmithKline clinical trials. *N Engl J Med*. 2013;369(5):475–8.
- Johnson & Johnson announces clinical trial data sharing agreement with Yale School of Medicine [<https://www.jnj.com/media-center/press-releases/johnson-johnson-announces-clinical-trial-data-sharing-agreement-with-yale-school-of-medicine>]. Accessed 8 Apr 2015.
- Bristol-Myers Squibb expands access to clinical trial data through collaboration with academic research institute [<http://news.bms.com/press-release/bristol-myers-squibb-expands-access-clinical-trial-data-through-collaboration-academic>]. Accessed 8 Apr 2015.
- Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9. W264.
- Thompson R, Bandera E, Burley V, Cade J, Forman D, Freudenheim J, Greenwood D, Jacobs D, Kalliecharan R, Kushi L, et al. Reproducibility of systematic literature reviews on food, nutrition, physical activity and endometrial cancer. *Public Health Nutr*. 2008;11(10):1006–14.
- Anonymous. Worldwide collaborative observational study and meta-analysis on allogenic leukocyte immunotherapy for recurrent spontaneous abortion. Recurrent Miscarriage Immunotherapy Trialists Group. [Erratum appears in *Am J Reprod Immunol* 1994 Oct;32(3):255]. *Am J Reprod Immunol*. 1994;32(2):55–72.
- Pladevall-Vila M, Delclos GL, Varas C, Guyer H, Bragues-Tarradellas J, Anglada-Arisa A. Controversy of oral contraceptives and risk of rheumatoid arthritis: meta-analysis of conflicting studies and review of conflicting meta-analyses with special emphasis on analysis of heterogeneity. *Am J Epidemiol*. 1996;144(1):1–14.
- Hopayan K, Mugford M. Conflicting conclusions from two systematic reviews of epidural steroid injections for sciatica: which evidence should general practitioners heed? *Br J Gen Pract*. 1999;49(438):57–61.
- Anonymous. Meta-analysis under scrutiny. *Lancet*. 1997;350(9079):675.
- Naylor CD. Meta-analysis and the meta-epidemiology of clinical research. *BMJ*. 1997;315(7109):617–9.
- Petticrew M, Kennedy SC. Detecting the effects of thromboprophylaxis: the case of the rogue reviews. *BMJ*. 1997;315(7109):665–8.
- Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ*. 2012;344:d7202.
- Fu R, Selph S, McDonagh M, Peterson K, Tiwari A, Chou R, Helfand M. Effectiveness and Harms of Recombinant Human Bone Morphogenetic Protein-2 in Spine Fusion: A Systematic Review and Meta-analysis. *Ann Intern Med*. 2013;158:890–902.
- Simmonds MC, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Rodgers MA, Stewart LA. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. *Ann Intern Med*. 2013;158(12):877–89.
- Fu XH, Li MN, Zheng GH, Le YQ, Wang L. Waste recombinant DNA: effectiveness of thermo-treatment to manage potential gene pollution. *Environ Pollut*. 2009;157(8-9):2536–41.
- Fu R, Selph S, McDonagh M, Peterson K, Tiwari A, Chou R, Helfand M. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med*. 2013;158(12):890–902.
- Krumholz HM, Ross JS. A model for dissemination and independent analysis of industry data. *JAMA*. 2011;306(14):1593–4.
- Medtronic—systematic reviews [<http://yoda.yale.edu/medtronic-systematic-reviews>]. Accessed 8 Apr 2015.
- Medtronic—available data [<http://yoda.yale.edu/medtronicrbmp-2>]. Accessed 8 Apr 2015.
- Glassman SD, Carreon LY, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, Dimar JR. RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. *Spine*. 2008;33(26):2843–9.
- Resnick D, Bozic KJ. Meta-analysis of trials of recombinant human bone morphogenetic protein-2: what should spine surgeons and their patients do with this information? *Ann Intern Med*. 2013;158(12):912–3.
- Riew KD, Carragee EJ. Commentary: despite reports of catastrophic complications, why recombinant human bone morphogenetic protein-2 should be available for use in anterior cervical spine surgery. *Spine J*. 2012;12(10):900–1.

- 36 Cahill KS, Chi JH, Day A, Claus EB. Prevalence, complications, and hospital charges associated with use of bone-morphogenetic proteins in spinal fusion procedures. *JAMA*. 2009;302(1):58–66.
- 37 Laine C, Guallar E, Mulrow C, Taichman DB, Cornell JE, Cotton D, Griswold ME, Localio AR, Meibohm AR, Stack CB, et al. Closing in on the truth about recombinant human bone morphogenetic protein-2: evidence synthesis, data sharing, peer review, and reproducible research. *Ann Intern Med*. 2013;158(12):916–8.
- 38 Hsu WK. Recombinant human bone morphogenetic protein-2 in spine surgery. *JBJS Rev*. 2014;2(6).
- 39 Carragee EJ, Hurwitz EL, Weiner BK. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. *The Spine Journal*. 2011;11(6):471–91.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

